Abstract

Disclosed is an audio signal time-scale modification which utilizes variable length synthesis for the improvement of output audio quality and reduced cross-correlation computations for the reduction of computation loads to a processor. An analysis window consisting of *N+Kmax* audio samples is selected from an input stream of audio samples and is shifted by the predetermined interval along an output stream to find optimal shift *Km*, which ensures best cross-correlation between Nov audio samples of the analysis window and last *Nov* audio samples of the output stream, and a particular value of *Nov_f* at which a coefficient of correlation between them is larger than a reference value or is the maximum one among a plurality of coefficients of correlation calculated with varying the value of *Nov*. The audio samples involved in the calculation of cross-correlation are down-selected by the predetermined ratio from *Nov* audio samples of the analysis window and last *Nov* audio samples of the output stream, respectively. The analysis window may also be shifted by the plurality of audio samples per one shift. The audio samples ranged $Km\sim(N+Km+Nm-Nov)^{th}$ in the analysis window is determined as an add frame. The existing last *Nov_f* audio samples of the output stream are replaced with new *Nov_f* audio samples obtained by weighting and adding the overlapped parts, i.e., the first *Nov_f* audio samples of the add frame and the last *Nov_f* audio samples of the output stream, while remaining part of the add frame is simply appended to the tail of the new *Nov_f* audio samples in the output stream.